تحليل مشكلة عدم تجانس التباين والمقارنة بين اختبار اوبراين واختبار اوبراين المعدل باستعمال المحاكاة

Analysis of the problem heterogeneity of variance and The Comparative between the O'Brien test and modified O'Brien test by using simulation

أ. د عباس لفته كنيهر alafta@uowasit.edu.iq

م.م احمد رزاق عبد رمضان

جامعة واسط – كلية الإدارة والاقتصاد

ملخص البحث : من أهم المشاكل التي تواجهنا في تصميم التجارب هي مشكلة عدم تجانس التباين والتي تأتي من العينات المختلفة التي تتبع مجتمعات بتباينات مختلفة أو ربما نتيجة تعرض مشاهدات تلك العينات إلى التلوث الذي يؤدي إلى الوصول إلى قرارات غير صحيحة عند اختبار الفرضيات وبالتالي الحصول على نتائج غير دقيقة ، ومن هنا جاء هدف الدراسة في استعمال اختبار اوبراين للكشف عن مشكلة عدم تجانس التباين وتحليلها، بالإضافة الى عرض طريقة ومقارنة نتائجها مع الطريقة الأصلية ولغرض التوصل الى الطريقة الأكفأ والأفضل فقد تم استعمال المحاكاة بطريقة مونت كارلو وتمت المقارنة بين الطرائق بالاعتماد على معيار قوة الاختبار .

ABSTRACT: The most important problems that face us the design of experiments is the problem of heterogeneity of variance which comes from the different samples that follows contrast the various communities or maybe as a result of Offers views of those samples to the pollution leads to access to incorrect decisions when testing hypotheses and therefore get inaccurate results, hence the purpose of the study is to use O'Brien test for the detection of the problem of lack of homogeneity of variance and analysis, in addition to displaying the way proposed and compared their results with the original method for the purpose of reaching a more efficient way and the best it has been a way to use Monte Carlo simulations and the comparison was made between the methods based on standard power of the test.

Keywords: Power of The Test, O'Brien Test, Homogeneity of variance

1-المقدمة والهدف:

يعد تحليل التباين احد اهم فروع علم الاحصاء ذات التطبيقات الواسعة أذ يقوم على عدة افتراضات اساسية وفي حالة عدم توفر احد هذه الافتراضات يؤدي الى نتائج غير صحيحة واهم هذه الافتراضات هي مشكلة عدم تجانس التباين والتي تأتي من اختلاف العينات التي تتبع مجتمعات بتباينات مختلفة أو ربما نتيجة تعرض مشاهدات تلك العينات إلى التلوث مما يؤدي

إلى قرارات غير دقيقة او مظللة عند اختبار الفرضيات اذ يرتفع مستوى المعنوية تلقائيًا ولذلك يجب أن تكون الاختلافات العشوائية داخل المعاملات على درجة كبيرة من التجانس وبالتالي تكون الاختلافات العشوائية متساوية بالنسبة الى العينات المختلفة مما يمكن معه الحصول على تباين مشترك للخطأ العشوائي لجميع العينات , ويوجد هناك العديد من الاختبارات للكشف عن هذه المشكلة ومنها اختبار اوبراين .

ان الهدف من هذه الدراسة هو المقارنة بين اختبار اوبراين لتجانس التباين واختبار اوبراين و تتم المقارنة باستعمال اسلوب المحاكاة وذلك بالاعتماد على معيار قوة الاختبار .

من اهم المؤلفات الأحصائيه والدراسات حول اختبار اوبراين لتجانس التباين منها دراسة (Martin and Games 1977) لجدول تحليل التباين لاختبارات تجانس التباين في ظل غياب التوزيع الطبيعي وعدم تساوي احجام العينات. ودراسة (O'Brien 1979) حول جدول تحليل التباين العام لاختبار تجانس التباين الحصين. وفي العام نفسه قام الباحثان (Games and Keselman) بدراسة اختبارات تجانس التباين وتأثيرها في التجارب العاملية . وقام الباحثان (Games and Keselman 1981) بدراسة بعنوان تأثيرات اختبار التباين البسيط في تصاميم التجارب الأحصائيه . واجرى العالم (Abdi 2007) دراسة لاختبار تجانس التباين اوبراين وقام بتطبيقه على عينتين متساوية بالحجم .كما قدم الباحثان (Lee and Katz and Restori 2010) دراسة حول مجموعة اختبارات تجانس التباين والمقارنة فيما بينهم باستعمال مونت كارلو . كما واجرى كل من الباحثان (Lemeshko and Gorbuneva 2012) دراسة لتطبيق الاختبارات المعلمية لتجانس التباين في ظل انتهاك الافتراضات الأساسية لتحليل التباين.

2- الجانب النظري

سيتم في هذا المبحث تحليل التباين من حيث الافتراضات والنموذج الرياضي وكذلك تناول اختبار اوبراين لتجانس التباين كما سيتضمن ايضا عرضا للطريقة المقترحة للاختبار .

2-1 تحليل التباين [4] [8]

أن تحليل التباين هو احد الأدوات الاحصائية المهمة والتي تعني بعملية دراسة العلاقة بين متغير تابع مع متغير اخر او عدة متغيرات مستقلة والتي عادة ما تكون وصفية ويركز اسلوب تحليل التباين في عمله على تحديد مصادر الأختلاف بين المتوسطات ولايهتم في تحديد نوع العلاقة بين المتغيرات التابعة والمتغيرات المستقلة ، وهناك عدة مستويات لتحليل التباين منها تحليل التباين بعامل واحد و تحليل التباين بعاملين و تحليل التباين للبيانات المتكررة ثم الى انواع اخرى لايتسع المجال لذكرها ، ويمكن صياغة الأنموذج الرياضي لتحليل التباين على النحو الأتي :

$$\gamma_{ij} = \mu_i + \varepsilon_{ij}$$

$$i = 1,2, \dots, r; j = 1, \dots, n_i$$

اذ ان γ_{ij} تمثل القيمة j للمتغير التابع في المستوى التصنيفي i ، وتمثل μ_i قيمة متوسط المتغير في المستوى التصنيفي i ، وتمثل ε_{ij} الخطأ العشوائي .

2-2 افتراضات تحليل التباين [1][3]

ان اهم شروط استعمال اسلوب تحليل التباين هي :

- ♦- ان تكون العينات مستقلة
- ♦- ان تكون العينات مسحوبة من مجتمعات ذات توزيع طبيعي
- ♦- أن يكون تباين المجتمعات متساوية أي ان التباينات تمثل بالفرضية الأتية :

$$\sigma_1{}^2 = \sigma_2{}^2 = \sigma_3{}^2 = \cdots = \sigma_k{}^2$$

2-3 تجانس التباينات [2][3]

للحصول على نتائج وقرارات صائبة هناك بعض الشروط يجب توافرها في البيانات، وعلى الباحث التأكد من توفر هذه الشروط قبل القيام بتحليلها، وأهم هذه الشروط هو تجانس التباينات والذي يعد شرطا اساسيا من شروط تحليل التباين والذي يتمثل بأن تكون المجتمعات التي سحبت منها المجموعات موضع المقارنة لها تباينات متساوية. ومن الجدير بالذكر هنا هو وجود عدة طرائق للكشف عن تجانس التباينات أو لاختبار الفرضية الآتية:

$$H_{0:}\ \sigma_1{}^2 = \sigma_2{}^2 = \sigma_3{}^2 = \cdots = \sigma_k{}^2$$

$$H_{1:}\ at\ least\ two\ of\ them\ are\ not\ equal$$

وهناك العديد من الاختبارات التي تكشف عن شرط تجانس التباين : سنتناول هنا احد اهم الاختبارات الأحصائيه للكشف عن هذه المشكلة وهو اختبار اوبراين واختبار اوبراين المقترح ولكن سنتطرق اولاً لقوة الاختبار وتعريفة لكونه يعتبر معيار المقارنة للمفاضلة بين الاختبارين.

2-4 قوة الاختبار [6][7] (Power Of the Test)

يمكن تعريف قوة الاختبار الإحصائي بأنه احتمال رفض فرضية العدم H_0 وهي خاطئة، والذي يعتمد على كل من الخطأ من النوع الأول (α) والخطأ من النوع الثاني (β) كما يعتمد على حجم العينة . وقوة الاختبار الاحصائي تساوي ($1-\beta$) ، أذ يمكن زيادة قوة الاختبار الإحصائي عن طريق مستوى الدلالة وحجم العينة . ويستعمل اغلب الإحصائيين مستويات الدلالة (0.01 ، 0.05) وهو أمر متفق علية وليس له دليل علمي ومنطقي . حيث ان قوة الاختبار ذات علاقة بالقيمة β ، وكلما زاد حجم β انخفض مقدار قوة الاختبار ان قوة الاختبار لأسلوب معين هو دالة في عدة متغيرات هي :-

- حجم العينة: حيث تزداد قوة الاختبار بزيادة حجم العينة.
- مستوى الدلالة المتوقع: حيث تزداد قوة الاختبار بارتفاع مستوى الدلالة.
- علاقة القيمة الحقيقية للمعلمة بقيمة الفرضية الصفرية: بافتراض ان الاختبار بذيلين تزداد قوة الاختبار كلما ابتعدت القيمة الحقيقية للمعلمة عن القيمة المفروضة، كذلك تزداد قوة الاختبار كثيرا بابتعاد القيمة الحقيقة عن القيمة المفروضة للجهتين الأعلى والقيم الأقل ، وتكون قوة الاختبار في نهايتها الصغرى عندما تكون القيمة الحقيقية مساوية للقيمة المفروضة .
- اذا كان الاختبار بذيل واحد او ذيلين: اذا كان الفرض البديل متجها كان الاختبار ذو نهاية واحدة (ذيل)، أما اذا كانا لفرض البديل عديم الاتجاه فأنه يسمى اختبار ذي

اتجاهين (ذيلين) وان اختبار الفرضية الصفرية بذيل واحد يزيد من قوة الاختبار الإحصائي . وبناء على مستوى الدلالة المعين من الباحث وبالاستفادة من توزيع المعاينة للمختبر الإحصائي يمكن ان تعين حدا حرجا للرفض أو قبول الفرضية الصفرية مع الأخذ بالحسبان الفرضية البديلة

اختبار اوبراين [5][7] [10] **O'Brien Test**

اختبار اوبراين هو أحد الاختبارات الأحصائيه التي تهتم بدراسة تجانس التباينات الذي هو أحد اهم الافتراضات الأساسية مع معظم الإجراءات الأحصائيه مثل تحليل التباين، ومن المهم أن تكون قادر على اختبار هذه الفرضية، حيث يعتبر اختبار اوبراين واحداً من الاختبارات الأكثر حساسية وقد تم تطوير هذا الاختبار من قبل اوبراين (1979، 1981). ان فرضية العدم لهذا الاختبار هي أن العينات تأتي من مجتمعات لها نفس التباين والفرضية البديلة هي أن يكون على الأقل اثنين من تباينات المجتمعات مختلفة:

$$H_{0:}\ {\sigma_1}^2 = {\sigma_2}^2 = {\sigma_3}^2 = \cdots = {\sigma_k}^2$$

$$H_{1:}\ at\ least\ two\ of\ them\ are\ not\ equal$$

أن ميزة اختبار اوبراين تكمن في تنوعه وانسجامه بالمقارنة مع غيره من اختبارات تجانس التباين مع تصاميم تحليل التباين المعياري، وايضا يعد الأفضل لأنه يقلل من الخطأ من النوع الأول والخطأ من النوع الثاني.

أن الفكرة الأساسية وراء اختبار اوبراين هي استبدال كل عينة، حيث يتم استبدال القيم الأصلية بالقيم المتحولة بحيث ان القيم المتحولة تعكس تباين العينة، ومن ثم فأن تحليل التباين القياسي يعتمد على القيم المتحولة وبالتالي فأن هذا التحليل يكشف عن افتراض تجانس التباين.

وأيضا من المميزات المهمة لاختبار اوبراين هي القدرة على استخدامه في حالة وجود عينتين وأكثر وكذلك قدرة استعماله مع مختلف تصاميم تحليل التباين الاحصائية.

بالنسبة الى تطبيق هذا الاختبار هناك تطبيق مباشر وهو أن تحل محل القيم الأصلية القيمة المطلقة لانحرافها إلى الوسط الحسابي للمجموعة الخاصة بهم. والتحويل يكون وفقاً للاتي:

$$u_{as} = |Y_{as} - M_a|$$

حيث ان :

Y_{as} : مشاهدات المجموعة a

M_a: الوسط الحسابي للمجموعة a

هذا التحويل يكون مفيد لكونه سهل الفهم وبسيط ولكن للأسف فانة يخلق بعض المشاكل الإحصائية منها أنه يؤدي الى زيادة حجم الخطأ من النوع الأول لذلك ان أفضل تطبيق لتحويل القيم الأصلية يكون من خلال استخراج القيمة المطلقة لانحراف المشاهدات عن وسيط المجموعة كما يأتي:

$$W_{as} = |Y_{as} - Md_a|$$

حيث ان :

Md_a : وسيط المجموعة a

وان هذا التحويل يعطي نتائج مرضية للغاية لاختبار افتراض تجانس التباين ولكن من اجل تنفيذ اجراءات احصائية أكثر تعقيداً مثل تحليل التباين والمقارنات المتعددة تم اقتراح طريقة تحويل اخرى من قبل العالم اوبراين (1981-1979) ويمكن استعمال هذه الطريقة عندما تكون احجام العينات غير متساوية وتتم طريقة تحويل القيم الأصلية كما هو واضح في القانون التالي:

$$u_{as} = \frac{N_a(N_a - 1.5)(Y_{as} - M_a)^2 - 0.5SS_a}{(N_a - 1)(N_a - 2)}$$

حيث ان:

N_a : عدد المشاهدات في المجموعه a

M_a: الوسط الحسابي للمجموعه a

$$SS_a = \sum_s (Y_{as} - M_a)^2$$

وعندما تكون احجام المجموعات او العينات متساوية فأن هذه الطريقة يمكن تبسيطها كما واضح ادناه:

$$u_{as} = \frac{N(N - 1.5)(Y_{as} - M_a)^2 - 0.5\ SS_a}{(N - 1)(N - 2)}$$

N: تمثل عدد المشاهدات في المجموعات الفرعية.

حيث ان :

$S_i{}^2$: يمثل تباين القيم الأصلية

u_{as}: يمثل القيم بعد التحويل

n_i : يمثل حجم العينة

احصاءه اختبار اوبراين تستخرج بحساب قيمة F التي تحسب في جدول ANOVA العادي باستعمال البيانات بعد التعديل او المتحولة، وتقارن قيمة F المحسوبة مع قيمة F الجدولية بدرجة حرية $F_{(k-1),(N-k)}$ وبمستوى معنوية (0.05) وتقبل فرضية العدم التي تنص على ان التباينات متجانسة إذا كانت قيمة $F_{tab>}F_{cal}$ وترفض إذا حصل العكس.

الطريقة المقترحة

في الطريقة الأصلية تم استعمال مجموع انحرافات مربعات القيم عن وسطها الحسابي ومن البديهي انها تساوي اقل ما يمكن: -

$$\sum (Y_{as} - M_a)^2 = min$$

أما في الطريقة المقترحة تم استعمال الحدود المطلقة واحد مقاييس النزعة المركزية (المنوال) لأنه يعطي اقل مجموع فروق للقيم المطلقة في حالة استعمال الوسط الحسابي وذلك لأنه يمكن تعريف المنوال على انه القيم الاكثر تكرارا في البيانات فعند اخذ مجموع الفروق المطلقة للقيم عن المنوال المساوي لأغلب قيم العينة فمن الطبيعي يكون الناتج اقل ما يمكن وبذلك نحصل على اقل مجموع فروق مطلقة مع الحفاظ او المحافظة على الفكرة الأصلية للاختبار والتي تبحث عن اقل مجموع فروق مطلقة او مربعة للبيانات ، وكمرحلة تمهيدية قام الباحث باحتساب الحد المطلق لانحرافات القيم عن وسطها الحسابي كما في القانون التالي :

$$u_{as} = \frac{N(N-1.5)|Y_{as} - M_a| - 0.5\sum_{i=1}^{ni}|Y_{as} - M_a|}{(N-1)(N-2)}$$

وفي المرحلة التالية تم باستبدال المنوال بدل الوسط الحسابي وكالاتي: -

$$\sum (Y_{as} - Mode)$$

وبذلك فأن مجموع انحرافات القيم المطلقة عن المنوال تكون اقل ما يمكن وبالتالي فان الطريقة المقترحة تكون أفضل في الكشف عن مشكلة تجانس تباين عينات المجتمع قيد الدراسة بعد الحصول على نتائج المحاكاة التي تعرض لاحقاً:

$$u_{as} = \frac{N(N-1.5)|Y_{as} - Mode| - 0.5\ \sum_{i=1}^{ni}|Y_{as} - mode|}{(N-1)(N-2)}$$

احصاءه اختبار اوبراين سوف تتم بحساب قيمة F التي تحسب في جدول ANOVA العادي باستعمال البيانات بعد التعديل او المتحولة، وتقارن قيمة Fالمحسوبة مع قيمة F الجدولية بدرجة حرية $F_{(k-1),(N-k)}$ وبمستوى معنوية (0.05) وتقبل فرضية العدم التي تنص على ان التباينات متجانسة إذا كانت قيمة $F_{tab>}F_{cal}$ وترفض اذا حصل العكس. وتمت عملية المقارنة بين الطريقتين بالاعتماد على معيار قوة الاختبار Power Of The Test.

3-الجانب التجريبي / وصف تجربة المحاكاة

لأجل الوصول الى الهدف المتوخى من الدراسة وهو المقارنة بين طرائق اختبار فرضية عدم تجانس التباين تم اللجوء الى استخدام المحاكاة لتوليد البيانات لغرض الوصول الى الطريقة المثلى للكشف عن مشكلة عدم تجانس التباين وقد تم استعمال اسلوب مونت كارلو.

باستعمال لغة البرمجة الماتلاب إذ تم توليد اربع عينات تتوزع توزيعاً كاوسياً (Gausian distribution) وكانت هذه العينات متساوية حيث تم اخذ المجتمع بحجم N وتقسيم المجتمع الى عينات متساوية بحيث ان $(\mathrm{n}_1 + n_2 + \mathrm{n}_3 + n_4 = N)$ مما يؤدي الى مستويات مختلفة من درجات التجانس وتم تغيير حجوم العينات اكثر من مرة لنرى ماهو تأثير التغير في حجم العينة على قوة الأختبار .

• تم افتراض قيم μ وقيم التباين σ^2 بحيث تتولد العينات بالاعتماد على القيم المفترضة وكانت هذا القيم قريبة جداً من بعضها بحيث يصعب على الاختبارات المستعملة الكشف عن مشكلة عدم تجانس التباين لنلاحظ دقة هذه الاختبارات في اصعب الظروف .

• تم توليد مجموعة من المتغيرات التي تتبع توزيعات مختلفة وهي :-

- التوزيع الطبيعي *(Normal dist.)*

- التوزيع الأسي *(Expo dist.)*

- توزيع كاما (*Gamma dist.*)

• تم تكرار التجربة 1000مرة وفي حالة كل من التوزيعات المذكورة، إذ تم افتراض مستويين للمعنوية هما 0.01 و 0.05 وتمت كتابة البرامج الخاصة بأسلوب المحاكاة Monte Carlo ، وكان مقياس المفاضلة بين اختبار اوبراين والطريقة المقترحة هو عن طريق معيار قوة الاختبار Power Of The Test ، وصيغة هذا المعيار هي :

$$\mathrm{P}(T) = \frac{K(T)}{R}$$

وان :

$K(T)$:هو عدد مرات رفض احصاءة الاختبار (T)للفرضية (H_0)الخاطئة

R :عدد مرات تكرار التجربة.

1-3 توليد بعض المتغيرات العشوائية المستمرة

توليد متغيرات عشوائية ذات توزيع طبيعي : هناك عدد من الطرائق لتوليد المتغيرات العشوائية ذات التوزيع الطبيعي وقد تم استعمال طريقة Box-Muller لتوليد متغيرات العشوائية مستمرة من التوزيع الطبيعي بمتوسط وتباين مفترضين مسبقاً وللحصول على المتغيرات يتم ذلك باستعمال الصيغة الأتية :

$$Z_1 = (-2\ln u_1)^{\frac{1}{2}}\ \cos 2\pi\, u_2$$

$$= (-2\ln u_1)^{\frac{1}{2}} Z_2\ \mathrm{Sin}2\pi\, u_2$$

$$Z_3 = (-2\ln u_3)^{\frac{1}{2}}\ \cos 2\pi\ u_4$$

$$= (-2\ln u_3)^{\frac{1}{2}} Z_4\ \mathrm{Sin}2\pi\ u_4$$

ثم تحويلها الى المتغيرات $(z_4 z_1, z_2, z_3,)$ إذ تم الحصول على متغيرات العشوائية التي تتبع التوزيع الطبيعي بمتوسط μ وتباين σ^2 يكون باستخدام الصيغة التالية:

$$Z = \frac{X_i - \mu}{\sigma} \text{ ، وبالتالي فأن } X_i = M + \sigma Z$$

$$X_1 = M + \sigma Z_1$$

$$X_2 = M + \sigma Z_2$$

$$X_3 = M + \sigma Z_3$$

$$X_4 = M + \sigma Z_4$$

وكانت قيم μ المفترضة تساوي (9،8،6،7) وقيم σ المفترضة تساوي (1.1،1.2،1.3،1.4) إذ تم توليد البيانات بالاعتماد على هذه القيم المفترضة.

توليد متغيرات عشوائية تتبع التوزيع الأسي

أن دالة الكثافة الاحتمالية للمتغير x العشوائي لهذا التوزيع هي:

$$F(X) = 1 - e^{-lx}$$

وان دالة التوزيع التجميعية c.d.f تأخذ الشكل الاتي:

$$U = 1 - e^{\lambda x} \qquad x \geq 0$$

$$e^{-\lambda x} = 1 - U$$

$$-\lambda x = \ln(1 - U)$$

$$X = -(1\ /\ \lambda)\ln\ (1 - U)$$

توليد متغيرات عشوائية تتبع توزيع كاما

ان دالة الكثافة الاحتمالية للمتغير x العشوائي لهذا التوزيع هي:

$$F(X) = \frac{1}{\Gamma(k) b^a} x^{a-1} e^{-x/b} \qquad , for\ \ x > 0\ and\ a, b > 0$$

وان دالة التوزيع التجميعية c.d.f تأخذ الشكل الاتي:

$$X = -b * \ln\big(2 * (1 - u)\big) + a$$

2-3 نتائج تحليل تجارب المحاكاة

1-التجربة الاولى (التوزيع الطبيعي)

نتائج اختبار عدم تجانس التباين الخاصة بالتوزيع الطبيعي عرضت في الجداول (2،1) والتي من خلالها تبين لنا في جدول (1) ان الطريقه التمهيدية كان هناك زيادة نسبية طفيفة في قوة الاختبار بالنسبة الطريقة الأصلية ويظهر ذلك كلما تكبر حجم العينة، هذا يعني ان هناك علاقة طردية بين حجم العينة وقوة الاختبار فكلما يكبر حجم العينة كلما تكبر قوة الاختبار.

جدول (1) يمثل نتائج قوة الاختبار في حالة التوزيع الطبيعي (المطلق والوسط الحسابي)

الاختبار / حجم العينة	F	اوبراين O	F	اوبراين المقترح*O
n=10	1.2803	0.4810	1.1529	0.5380
n=30	1.0356	0.5660	1.5013	0.6270
n=60	2.0054	0.7130	2.5086	0.7470
n=100	1.9842	0.8190	2.1922	0.8590

اما بالنسبة للطريقة المقترحة الموضحة في جدول (2) نلاحظ ان هناك افضلية واضحة لهذه الطريقة بالمقارنة مع طريقة اوبراين الأصلية ويرجع سبب هذه الأفضلية في وجود أحد مقاييس النزعة المركزية وهو المنوال الذي يؤدي الي تقليل انحرافات القيم المطلقة عن المنوال وبهذا نكون قللنا من قيمة الخطأ من النوع الاول وبالتالي فان الطريقة المقترحة تكون أفضل في الكشف عن مشكلة تجانس تباين عينات المجتمع قيد الدراسة.

ونلاحظ ان هذه الطريقة تعطي افضلية للعينات الكبيرة أكثر من العينات الصغيرة كما هو موضح في الجدول الاتي: -

$\pi\ (O^*)$	N
1	60
1	100

جدول (2) يمثل نتائج قوة الاختبار في حالة التوزيع الطبيعي (المطلق والمنوال)

الاختبار / حجم العينة	F	اوبراين O	F	اوبراين المقترح*O
n=10	0.6920	0.4770	7.0076	0.7860
n=30	3.0776	0.5620	39.4246	0.9950
n=60	3.5563	0.7040	56.8028	1
n=100	4.5772	0.8160	60.0916	1

2- <u>التجربة الثانية (التوزيع الأسي)</u>

نتائج اختبار عدم تجانس التباين الخاصة بالتوزيع الأسي عرضت في الجداول (4،3) والتي من خلالها تبين لنا في جدول (3) ان الطريقه التمهيدية كانت هناك زيادة واضحة في قوة الاختبار بالنسبة

الطريقة الأصلية ويظهر ذلك كلما تكبر حجم العينة ، هذا يعني ان هناك علاقة عكسية بين حجم العينة وقوة الاختبار فكلما يكبر حجم العينة كلما تكبر قوة الاختبار .

جدول (3) يمثل نتائج قوة الاختبار في حالة التوزيع الأسي (المطلق والوسط الحسابي)

الاختبار / حجم العينة	F	اوبراين O	F	اوبراين المقترح *O
n=10	0.2351	0.5210	0.4790	0.7170
n=30	0.2120	0.5070	0.7200	0.7330
n=60	1.5158	0.5780	2.7494	0.7760
n=100	0.8794	0.6380	2.2803	0.8370

اما بالنسبة للطريقة المقترحة الموضحة في جدول (3) نلاحظ ان هناك افضلية واضحة لهذه الطريقة بالمقارنة مع طريقة اوبراين الأصلية ويرجع سبب هذه الأفضلية في وجود أحد مقاييس النزعة المركزية وهو المنوال الذي يؤدي الي تقليل انحرافات القيم المطلقة عن المنوال وبهذا نكون قللنا من قيمة الخطأ من النوع الاول وبالتالي فان الطريقة المقترحة تكون افضل في الكشف عن مشكلة تجانس تباين عينات المجتمع قيد الدراسة. ولكن قوة الاختبار للطريقة المقترحة تكون اقل من الطريقة الاصلية في حالة حجوم العينات الصغيرة.

وان اعلى قوة اختبار تم الحصول عليها في حالة التوزيع الاسي موضحة في الجدول الاتي: -

$\pi\ (O^*)$	N
0.6370	60
0.7370	100

جدول (4) يمثل نتائج قوة الاختبار في حالة التوزيع الأسي (المطلق والمنوال)

الاختبار / حجم العينة	F	اوبراين O	F	اوبراين المقترح*O
n=10	0.5600	0.5090	1.8804	0.3920
n=30	2.6692	0.5120	1.8791	0.5050
n=60	0.2241	0.5830	0.5553	0.6370
n=100	3.7460	0.6800	3.9166	0.7370

3- التجربة الثالثة (توزيع كاما)

نتائج اختبار عدم تجانس التباين الخاصة بتوزيع كاما عرضت في الجداول (5،6) والتي من خلالها تبين لنا في جدول (5) ان الطريقه التمهيدية كانت لها قوة الاختبار اعلى مقارنة مع الطريقة

الأصلية ويظهر ذلك كلما تكبر حجم العينة ، هذا يعني ان هناك علاقة عكسية بين حجم العينة وقوة الاختبار فكلما يكبر حجم العينة كلما تكبر قوة الاختبار .

جدول (5) يمثل نتائج قوة الاختبار في حالة توزيع كاما (المطلق والوسط الحسابي)

الاختبار / حجم العينة	F	اوبراين O	F	اوبراين المقترح*O
n=10	0.8704	0.5210	1.0258	0.7330
n=30	0.5676	0.5410	1.7315	0.7190
n=60	1.3733	0.5840	2.0042	0.7870
n=100	2.9592	0.6560	4.9098	0.8470

اما بالنسبة للطريقة المقترحة الموضحة في جدول (6) نلاحظ هناك افضلية واضحة لهذه الطريقة بالمقارنة مع طريقة اوبراين الأصلية ويرجع سبب هذه الأفضلية في وجود احد مقاييس النزعة المركزية وهو المنوال الذي يؤدي الي تقليل انحرافات القيم المطلقة عن المنوال وبهذا نكون قللنا من قيمة الخطأ من النوع الاول وبالتالي فان الطريقة المقترحة تكون افضل في الكشف عن مشكلة تجانس تباين عينات المجتمع قيد الدراسة. يمكن ان نرى من خلال النتائج ان قوة اختبار الطريقة المقترحة في حالة توزيع كاما تكون ضعيفة في حالة العينات الصغيرة فمثلا عندما حجم العينة 10 فان قوة الاختبار تساوي 0.4180 وهذه قيمة قليلة مقارنة مع الطريقة الأصلية.

وان اعلى قوة اختبار تم الحصول عليها في حالة توزيع كاما موضحة في الجدول الاتي: -

$\pi\ (O^*)$	N
0.6430	60
0.7630	100

جدول (6) يمثل نتائج قوة الاختبار في حالة توزيع كاما التعديل الثاني (المطلق والمنوال)

الاختبار / حجم العينة	F	اوبراين O	F	اوبراين المقترح*O
n=10	2.3078	0.5010	0.7292	0.4180
n=30	2.1363	0.5150	1.5687	0.5310
n=60	1.0739	0.5820	0.4797	0.6430
n=100	1.1009	0.6290	1.8884	0.7630

الاستنتاجات

بالاعتماد على تحليل النتائج لأسلوب المحاكاة تم التوصل الى الاستنتاجات الاتية: -

1- من خلال احتساب قوة الاختبار للطريقتين الأصلية والمقترحة ولثلاث من التوزيعات الأحصائيه وهي (التوزيع الطبيعي، التوزيع الاسي، توزيع كاما) فأننا نستنتج بان هناك علاقة طردية بين قوة الاختبار وحجم العينة فكلما يكبر حجم العينة كلما تكبر قوة الاختبار.
2- من خلال النظر الى نتائج قوة الاختبار نلاحظ افضلية الطريقة المقترحة وتفوقها على الطريقة الأصلية.
3- بعد استعمال ثلاث من التوزيعات وهي (التوزيع الطبيعي، التوزيع الاسي، توزيع كاما) نلاحظ ان اعلى قوة اختبار للطريقة المقترحة تم الحصول عليها كانت في حالة التوزيع الطبيعي فقد كانت مساوية الى الواحد في حالة العينات الكبيرة.
4- لقد تفوقت الطريقة المقترحة بسبب استعمال أحد مقاييس النزعة المركزية وهو المنوال الذي يقلل من انحرافات القيم المطلقة عن المنوال وبالتالي فأن الطريقة المقترحة تكون أفضل في الكشف عن مشكلة تجانس تباين عينات المجتمع قيد الدراسة.
5- لا يمكن استخدام الطريقة المقترحة في حالة وجود بيانات لا تحتوي على منوال لذا تستعمل الطريقة الاصلية لاختبار اوبراين.

التوصيات: -

1- عند الكشف عن مشكلة تجانس التباينات يفضل استخدام الطريقة المقترحة.
2- عند استخدام اختبار اوبراين للتجانس التباين يفضل ان تتوزع العينات توزيعاً طبيعيا لأن قوة الاختبار تكون اعلى في التوزيع الطبيعي.
3- نوصي باستعمال اختبار اوبراين لتجانس التباين في حالة وجود عينات كبيرة.
4- اخذ توزيعات احتمالية جديدة للبيانات ودراستها.
5- عمل دراسة لاختبار اوبراين في حالة وجود قيم شاذة (*outliers*) في البيانات

المصـــادر:

1- العتيبـــي ، أشـــرف احمـــد عـــواض ، (2011)، " دراســـة تقويميـــة لصـــحة اســـتخدام أسـلوب تحليـل التبـاين فـي رسـائل الماجسـتير والـدكتوراه فـي كليـة التربيـة فـي جامعـة أم القـرى (عبـر الفتـرة الزمنيـة ١٤٢١ ه - ١٤٣٠) "، رسـالة ماجسـتير فـي بحـوث العمليات ، كلية التربية ،قسم علم النفس ، جامعة ام القرى .

2- الشـمراني ، محمـد موسـى محمـد ، (2000) ،" مشـكلات اسـتخدام تحليـل التبـاين الأحـادي والمقارنـات البعدية وطـرق علاجهـا " ، رسـالة ماجسـتير فـي بحـوث العمليات ،كلية التربية ، جامعة ام القرى ، مكة .

3- المشـهداني كمـال علـوان ، عمـاد حـازم ،(2009)، "اختبـار الفرضـيات الأحصـائيه " ، الجزيرة للطباعة والنشر /جامعة بغداد .

4- المشـهداني ، كمـال علـوان والبيـاتي ،محمـود مهـدي ،(2014)،" التصـاميم التجريبيـة وتحليـل التجـارب باسـتعمال برنـامج spss " ، كليـة الادارة والاقتصـاد ، جامعـة بغداد.

5- Lee H.B., Katz G.S. and Restori A.F(2010),"A Monte carlo Study of seven homogeneity of variance tests", Journal of Mathematics and Statistics, vol.6(3) , pp. 359-366 .

6- Lemeshko B.Yu. and Lemeshko S.B.(2008)," Power and robustness of criteria used to verify the homogeneity of means", Measurement Techniques, 51, 9, pp 950–959.

7- Lemeshko,B.Y.andGorbuneva,A.A.(2012),"Appllication of paramet-eric homogeneity of variance test under violation of classical assumption ",proceedings , 2 nd stochastic modeling techniques and data analysis international conferece 5-8 , Chaina Crete Greece.

8- Wiley J. and Sons. (2005), " Design and Analysis Of Experiments", Hoboken, New Jersey and Canada.

9- William F.G.(2003)," Lectures on Statistics "

المصادر الألكترونية :

10 - https://www.utdallas.edu/~herve/Abdi-OBrien2007-pretty.pdf

البرنامج

```
function [OBrientest] = OBrientest(X,alpha)
if nargin < 2;
   alpha = 0.05;
end
clc;
n1=100;
n2=100;
n3=100;
n4=100;
r=1000;
m1=8;
m2=7;
m3=5;
m4=6;
sigma1=1.4;
sigma2=1.3;
sigma3=1.2;
sigma4=1.1;
pi=22/7;
reject1=0;reject2=0;
lower=0.05-2*(0.05*(1-0.05)/10000)^(1/2);
upper=0.05+2*(0.05*(1-0.05)/10000)^(1/2);
for j=1:r;
u1=rand(n1,1);
u2=rand(n2,1);
u3=rand(n3,1);
u4=rand(n4,1);
zz1=(-2*log(u1)).^(1/2).*cos(2*pi*u2);
zz2=(-2*log(u1)).^(1/2).*sin(2*pi*u2);
zz3=(-2*log(u3)).^(1/2).*sin(2*pi*u4);
zz4=(-2*log(u4)).^(1/2).*cos(2*pi*u4);
y1=sigma1*zz1+m1;
y2=sigma2*zz2+m2;
y3=sigma3*zz3+m3;
y4=sigma4*zz4+m4;
for j=1:r;
end
mm=100;
v=(1:4);
o=repmat(v,mm,1);
o=[o(:)];
h=[y1;y2;y3;y4];
X=[h o];
Y=(X);
k=max(Y(:,2));
fprintf('The number of samples are:%2i\n\n', k);
%O'Brien Procedure.
n=[];s2=[];Z=[];
```

```
indice=Y(:,2);
for i=1:k
   Ye=find(indice==i);
   eval(['Y' num2str(i) '=Y(Ye,1);']);
   eval(['mY' num2str(i) '=mean(Y(Ye,1));']);
   eval(['n' num2str(i) '=length(Y' num2str(i) ') ;'])
   eval(['s2' num2str(i) '=(std(Y' num2str(i) ').^2) ;'])
   eval(['xn= n' num2str(i) ';'])
   eval(['xs2= s2' num2str(i) ';'])
   eval(['Z' num2str(i) '= ((n' num2str(i) ' - 1.5)*(n'
num2str(i) ')*((Y' num2str(i) ' - mY' num2str(i) ').^2)-
((0.5)*(s2' num2str(i) ')*(n' num2str(i) ' - 1)))/((n'
num2str(i) ' - 1)*(n' num2str(i) ' - 2));']);
   eval(['x= Z' num2str(i) ';']);
   n=[n;xn];s2=[s2;xs2];Z=[Z;x];
end
for i=1:k
   if n(i)==2
      error('Requires sample sizes greater than two.
Please, redefine the data matrix.');
   end
end
Y=[Z Y(:,2)];
fprintf('-----------------------------\n');
disp(' Sample    Size      Variance')
fprintf('-----------------------------\n');
for i=1:k
   fprintf('   %d       %2i         %.4f\n',i,n(i),s2(i))
end
fprintf('-----------------------------\n');
disp(' ')
C=(sum(Y(:,1)))^2/length(Y(:,1)); %correction term.
SST=sum(Y(:,1).^2)-C; %total sum of squares.
dfT=length(Y(:,1))-1; %total degrees of freedom.
indice=Y(:,2);
for i=1:k
   Ye=find(indice==i);
   eval(['A' num2str(i) '=Y(Ye,1);']);
end
A=[];
for i=1:k
   eval(['x =((sum(A' num2str(i) ').^2)/length(A'
num2str(i) '));']);
   A=[A,x];
end
SSA=sum(A)-C; %sample sum of squares.
dfA=k-1; %sample degrees of freedom.
SSE=SST-SSA; %error sum of squares.
dfE=dfT-dfA; %error degrees of freedom.
MSA=SSA/dfA; %sample mean squares.
```

```
MSE=SSE/dfE; %error mean squares.
F=MSA/MSE; %F-statistic.
v1=dfA;df1=v1;
v2=dfE;df2=v2;
P = 1 - fcdf(F,v1,v2);  %probability associated to the F-
statistic.
fprintf('O''Brien''s Test for Equality of Variances
F=%3.4f, df1=%2i, df2=%2i\n', F,df1,df2);
fprintf('Probability associated to the F statistic =
%3.4f\n', P);
if P >= alpha;
  fprintf('The associated probability for the F test is
equal or larger than% 3.2f\n', alpha);
  fprintf('So, the assumption of homoscedasticity was
met.\n');
else
  fprintf('The associated probability for the F test is
smaller than% 3.2f\n', alpha);
  fprintf('So, the assumption of homoscedasticity was not
met.\n');
end
if F>(P>=alpha)
    reject1=reject1+1;
end
% الطريقة
المقترحة$%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
mm=100;
v=(1:4);
o=repmat(v,mm,1);
o=[o(:)];
h=[y1;y2;y3;y4];
X=[h o];
Y=X;
k=max(Y(:,2));
fprintf('The number of samples are:%2i\n\n', k);
n=[];s2=[];Z=[];
indice=Y(:,2);
for i=1:k
   Ye=find(indice==i);
   eval(['Y' num2str(i) '=Y(Ye,1);']);
   eval(['mY' num2str(i) '=mode(Y(Ye,1));']);
   eval(['n' num2str(i) '=length(Y' num2str(i) ') ;'])
   eval(['s2' num2str(i) '=(std(Y' num2str(i) ').^2) ;'])
   eval(['xn= n' num2str(i) ';'])
   eval(['xs2= s2' num2str(i) ';'])
   eval(['Z' num2str(i) '= ((n' num2str(i) ' - 1.5)*(n'
num2str(i) ')*(abs(Y' num2str(i) ' - mY' num2str(i) '))-
((0.5)*(abs(Y' num2str(i) ' - mY' num2str(i) '))*(n'
```

```
num2str(i) ' - 1)))/((n' num2str(i) ' - 1)*(n' num2str(i)
' - 2));']);
   eval(['x= Z' num2str(i) ';']);
   n=[n;xn];s2=[s2;xs2];Z=[Z;x];
end
for i=1:k
   if n(i)==2
      error('Requires sample sizes greater than two.
Please, redefine the data matrix.');
   end
end
Y=[Z Y(:,2)];
fprintf('-----------------------------\n');
disp(' Sample    Size      Variance')
fprintf('-----------------------------\n');
for i=1:k
   fprintf('   %d       %2i          %.4f\n',i,n(i),s2(i))
end
fprintf('-----------------------------\n');
disp(' ')
C=(sum(Y(:,1)))^2/length(Y(:,1)); %correction term.
SST=sum(Y(:,1).^2)-C; %total sum of squares.
dfT=length(Y(:,1))-1; %total degrees of freedom.
indice=Y(:,2);
for i=1:k
   Ye=find(indice==i);
   eval(['A' num2str(i) '=Y(Ye,1);']);
end
A=[];
for i=1:k
   eval(['x =((sum(A' num2str(i) ').^2)/length(A'
num2str(i) '));']);
   A=[A,x];
end
SSA=sum(A)-C; %sample sum of squares.
dfA=k-1; %sample degrees of freedom.
SSE=SST-SSA; %error sum of squares.
dfE=dfT-dfA; %error degrees of freedom.
MSA=SSA/dfA; %sample mean squares.
MSE=SSE/dfE; %error mean squares.
F=MSA/MSE; %F-statistic.
v1=dfA;df1=v1;
v2=dfE;df2=v2;
P = 1 - fcdf(F,v1,v2);  %probability associated to the F-
statistic.
fprintf('O''Brien''s Test for Equality of Variances
F=%3.4f, df1=%2i, df2=%2i\n', F,df1,df2);
fprintf('Probability associated to the F statistic =
%3.4f\n', P);
if P >= alpha;
```

```
  fprintf('The associated probability for the F test is
equal or larger than% 3.2f\n', alpha);
  fprintf('So, the assumption of homoscedasticity was
met.\n');
else
  fprintf('The associated probability for the F test is
smaller than% 3.2f\n', alpha);
  fprintf('So, the assumption of homoscedasticity was not
met.\n');
end
if F>(P>=alpha)
    reject2=reject2+1;
end
end
alfa1=reject1/r
alfa2=reject2/r
[lower upper]
end
```